GLOBAL EDITION

# Econometric Analysis

EIGHTH EDITION

William H. Greene

EIGHTH EDITION
GLOBAL EDITION

# ECONOMETRIC ANALYSIS

**William H. Greene**

*The Stern School of Business*

*New York University*

Pearson

Harlow, England • London • New York • Boston • San Francisco • Toronto • Sydney • Dubai • Singapore • Hong Kong
Tokyo • Seoul • Taipei • New Delhi • Cape Town • São Paulo • Mexico City • Madrid • Amsterdam • Munich • Paris • Milan

### *Example 10.6 Identification of a Supply and Demand Model*

Consider a market in which $q$ is quantity of $Q$, $p$ is price, and $z$ is the price of $Z$, a related good. We assume that $z$ enters both the supply and demand equations. For example, $Z$ might be a crop that is purchased by consumers and that will be grown by farmers instead of $Q$ if its price rises enough relative to $p$. Thus, we would expect $\alpha_2 > 0$ and $\beta_2 < 0$. So,

$$\begin{aligned} q_d &= \alpha_0 + \alpha_1 p + \alpha_2 z + \varepsilon_d \quad &&(\textit{demand}),\\ q_s &= \beta_0 + \beta_1 p + \beta_2 z + \varepsilon_s \quad &&(\textit{supply}),\\ q_d &= q_s = q \quad &&(\textit{equilibrium}). \end{aligned}$$

The reduced form is

$$q = \frac{\alpha_1\beta_0 - \alpha_0\beta_1}{\alpha_1 - \beta_1} + \frac{\alpha_1\beta_2 - \alpha_2\beta_1}{\alpha_1 - \beta_1} z + \frac{\alpha_1\varepsilon_s - \beta_1\varepsilon_d}{\alpha_1 - \beta_1} = \pi_{11} + \pi_{21} z + \nu_q,$$

$$p = \frac{\beta_0 - \alpha_0}{\alpha_1 - \beta_1} + \frac{\beta_2 - \alpha_2}{\alpha_1 - \beta_1} z + \frac{\varepsilon_s - \varepsilon_d}{\alpha_1 - \beta_1} = \pi_{12} + \pi_{22} z + \nu_p.$$

With only four reduced-form coefficients and six structural parameters, that there will not be a complete solution for all six structural parameters in terms of the four reduced form parameters. This model is unidentified. There is insufficient information in the sample and the theory to deduce the structural parameters.

Suppose, though, that it is known that $\beta_2 = 0$ (farmers do not substitute the alternative crop for this one). Then the solution for $\beta_1$ is $\pi_{21}/\pi_{22}$. After a bit of manipulation, we also obtain $\beta_0 = \pi_{11} - \pi_{12}\pi_{21}/\pi_{22}$. The exclusion restriction identifies the supply parameters; $\beta_2 = 0$ excludes $z$ from the supply equation. But this step is as far as we can go. With this restriction, the model becomes partially identified. Some, but not all, of the parameters can be estimated.

Now, suppose that income $x$, rather than $z$, appears in the demand equation. The revised model is

$$\begin{aligned} q &= \alpha_0 + \alpha_1 p + \alpha_2 x + \varepsilon_1,\\ q &= \beta_0 + \beta_1 p + \beta_2 z + \varepsilon_2. \end{aligned}$$

Note that one variable is now excluded from each equation. The structure is now

$$[q \quad p]\begin{bmatrix} 1 & 1 \\ -\alpha_1 & -\beta_1 \end{bmatrix} + [1\ x\ z]\begin{bmatrix} -\alpha_0 & -\beta_0 \\ -\alpha_2 & 0 \\ 0 & -\beta_2 \end{bmatrix} = [\varepsilon_1 \quad \varepsilon_2].$$

The reduced form is

$$[q \quad p] = [1\ x\ z]\begin{bmatrix} (\alpha_1\beta_0 - \alpha_0\beta_1)/\Lambda & (\beta_0 - \alpha_0)/\Lambda \\ -\alpha_2\beta_1/\Lambda & -\alpha_2/\Lambda \\ \alpha_1\beta_2/\Lambda & \beta_2/\Lambda \end{bmatrix} + [\nu_1 \quad \nu_2],$$

where $\Lambda = (\alpha_1 - \beta_1)$. The unique solutions for the structural parameters in terms of the reduced-form parameters are now

$$\begin{aligned} \alpha_0 &= \pi_{11} - \pi_{12}\left(\frac{\pi_{31}}{\pi_{32}}\right), \qquad & \beta_0 &= \pi_{11} - \pi_{12}\left(\frac{\pi_{21}}{\pi_{22}}\right),\\ \alpha_1 &= \frac{\pi_{31}}{\pi_{32}}, & \beta_1 &= \frac{\pi_{21}}{\pi_{22}},\\ \alpha_2 &= \pi_{22}\left(\frac{\pi_{21}}{\pi_{22}} - \frac{\pi_{31}}{\pi_{32}}\right), & \beta_2 &= \pi_{32}\left(\frac{\pi_{31}}{\pi_{32}} - \frac{\pi_{21}}{\pi_{22}}\right). \end{aligned}$$

With this formulation, all of the parameters are identified. This is an example of an exactly identified model. An additional variation is worth a look. Suppose that a second variable, $w$ (weather), appears in the supply equation,

$$q = \alpha_0 + \alpha_1 p + \alpha_2 x + \varepsilon_1,$$
$$q = \beta_0 + \beta_1 p + \beta_2 z + \beta_3 w + \varepsilon_2.$$

You can easily verify that, the reduced form matrix is the same as the previous one, save for an additional row that contains $[\alpha_1\beta_3/\Lambda, \beta_3/\Lambda]$. This implies that there is now a second solution for $\alpha_1$, $\pi_{41}/\pi_{42}$. The two solutions, this and $\pi_{31}/\pi_{32}$, will be different. This model is overidentified. There is more information in the sample and theory than is needed to deduce the structural parameters.

Some equation systems are identified and others are not. The formal mathematical conditions under which an equation in a system is identified turns on two results known as the rank and order conditions. The *order condition* is a simple counting rule. It requires that the number of exogenous variables that appear elsewhere in the equation system must be at least as large as the number of endogenous variables in the equation. (Other specific restrictions on the parameters will be included in this count—note that an "exclusion restriction" is a type of linear restriction.) We used this rule when we constructed the IV estimator in Chapter 8. In that setting, we required our model to be at least *identified* by requiring that the number of instrumental variables not contained in $\mathbf{X}$ be at least as large as the number of endogenous variables. The correspondence of that single equation application with the condition defined here is that the rest of the equation system is the source of the instrumental variables. One simple order condition for identification of an equation system is that each equation contain "its own" exogenous variable that does not appear elsewhere in the system.

The **order condition** is necessary for identification; the **rank condition** is sufficient. The equation system in (10-37) in structural form is $\mathbf{y}'\mathbf{\Gamma} = -\mathbf{x}'\mathbf{B} + \boldsymbol{\varepsilon}'$. The reduced form is $\mathbf{y}' = \mathbf{x}'(-\mathbf{B}\,\mathbf{\Gamma}^{-1}) + \boldsymbol{\varepsilon}'\mathbf{\Gamma}^{-1} = \mathbf{x}'\mathbf{\Pi} + \mathbf{v}'$. The way we are going to deduce the parameters in $(\mathbf{\Gamma}, \mathbf{B}, \mathbf{\Sigma})$ is from the reduced form parameters $(\mathbf{\Pi}, \mathbf{\Omega})$. For the $j$th equation, the solution is contained in $\mathbf{\Pi\Gamma}_j = -\mathbf{B}_j$, where $\mathbf{\Gamma}_j$ contains all the coefficients in the $j$th equation that multiply endogenous variables. One of these coefficients will equal one, usually some will equal zero, and the remainder are the nonzero coefficients on endogenous variables in the equation, $\mathbf{Y}_j$ [these are denoted $\gamma_j$ in (10-41) following]. Likewise, $\mathbf{B}_j$ contains the coefficients in equation $j$ on all exogenous variables in the model—some of these will be zero and the remainder will multiply variables in $\mathbf{X}_j$, the exogenous variables that appear in this equation [these are denoted $\boldsymbol{\beta}_j$ in (10-41) following]. The empirical counterpart will be $\mathbf{Pc}_j = \mathbf{b}_j$, where $\mathbf{P}$ is the estimated reduced form, $(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$, and $\mathbf{c}_j$ and $\mathbf{b}_j$ will be the estimates of the $j$th columns of $\mathbf{\Gamma}$ and $\mathbf{B}$. The rank condition ensures that there is a solution to this set of equations. In practical terms, the rank condition is difficult to establish in large equation systems. Practitioners typically take it as a given. In small systems, such as the two-equation systems that dominate contemporary research, it is trivial, as we examine in the next example. We have already used the rank condition in Chapter 8, where it played a role in the relevance condition for instrumental variable estimation. In particular, note after the statement of the assumptions for instrumental variable estimation, we assumed $\text{plim}(1/T)\mathbf{Z}'\mathbf{X}$ is a matrix with rank $K$. (This condition is often labeled the *rank condition* in contemporary applications. It not identical, but it is sufficient for the condition mentioned here.)

### *Example 10.7 The Rank Condition and a Two-Equation Model*

The following two-equation recursive model provides what is arguably the platform for much of contemporary econometric analysis. The main equation of interest is

$$y = \gamma f + \beta x + \varepsilon.$$

The variable $f$ is endogenous (it is correlated with $\varepsilon$.); x is exogenous (it is uncorrelated with $\varepsilon$). The analyst has in hand an instrument for $f$, $z$. The instrument, $z$, is relevant, in that in the auxiliary equation,

$$f = \lambda x + \delta z + w,$$

$\delta$ is not zero. The exogeneity assumption is $E[\varepsilon z] = E[wz] = 0$. Note that the source of the endogeneity of $f$ is the assumed correlation of $w$ and $\varepsilon$. For purposes of the exercise, assume that $E[xz] = 0$ and the data satisfy $\mathbf{x}'\mathbf{z} = 0$—this actually loses no generality. In this two-equation model, the second equation is already in reduced form; $x$ and $z$ are both exogenous. It follows that $\lambda$ and $\delta$ are estimable by least squares. The estimating equations for $(\gamma, \beta)$ are

$$\mathbf{P}\gamma_1 = \begin{bmatrix} \mathbf{x}'\mathbf{x} & \mathbf{x}'\mathbf{z} \\ \mathbf{z}'\mathbf{x} & \mathbf{z}'\mathbf{z} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{x}'\mathbf{y} & \mathbf{x}'\mathbf{f} \\ \mathbf{z}'\mathbf{y} & \mathbf{z}'\mathbf{f} \end{bmatrix} \begin{pmatrix} 1 \\ -\gamma \end{pmatrix} = \begin{bmatrix} \mathbf{x}'\mathbf{y}/\mathbf{x}'\mathbf{x} & \mathbf{x}'\mathbf{f}/\mathbf{x}'\mathbf{x} \\ \mathbf{z}'\mathbf{y}/\mathbf{z}'\mathbf{z} & \mathbf{z}'\mathbf{f}/\mathbf{z}'\mathbf{z} \end{bmatrix} \begin{pmatrix} 1 \\ -\gamma \end{pmatrix} = \boldsymbol{\beta}_{\mathrm{j}} = \begin{pmatrix} \beta \\ 0 \end{pmatrix}.$$

The solutions are $\gamma = (\mathbf{z}'\mathbf{y}/\mathbf{z}'\mathbf{f})$ and $\beta = (\mathbf{x}'\mathbf{y}/\mathbf{x}'\mathbf{x} - (\mathbf{z}'\mathbf{y}/\mathbf{z}'\mathbf{f})\mathbf{x}'\mathbf{f}/\mathbf{x}'\mathbf{x})$. Because $\mathbf{x}'\mathbf{x}$ cannot equal zero, the solution depends on $(\mathbf{z}'\mathbf{f}/\mathbf{z}'\mathbf{z})$ not equal to zero—formally that this part of the reduced form coefficient matrix have rank $M = 1$, which would be the rank condition. Note that the solution for $\gamma$ is the instrumental variable estimator, with $z$ as instrument for $f$. (The simplicity of this solution turns on the assumption that $\mathbf{x}'\mathbf{z} = 0$. The algebra gets a bit more complicated without it, but the conclusion is the same.)

The rank condition is based on the exclusion restrictions in the model—whether the exclusion restrictions provide enough information to identify the coefficients in the $j$th equation. Formally, the idea can be developed thusly. With the $j$th equation written as in (10-41), we call $\mathbf{X}_j$ the *included exogenous variables*. The remaining excluded exogenous variables are denoted $\mathbf{X}_j^*$. The $M_j$ variables $\mathbf{Y}_j$ in (10-41) are the included endogenous variables. With this distinction, we can write the $M_j$ reduced forms for $\mathbf{Y}_j$ as $\overline{\mathbf{\Pi}}_j = \begin{bmatrix} \mathbf{\Pi}_j \\ \mathbf{\Pi}_j^* \end{bmatrix}$. The rank condition (which we state without proof) is that the rank of the lower part of the $M_j \times (K_j + K_j^*)$ matrix, $\overline{\mathbf{\Pi}}_j$, equal $M_j$. In the preceding example, in the first equation, $\mathbf{Y}_j$ is $f$, $M_j = 1$, $\mathbf{X}_j$ is $x$, $\mathbf{X}_j^*$ is $z$, and $\overline{\mathbf{\Pi}}_j$ is estimated by the regression of $f$ on $x$ and $z$; $\mathbf{\Pi}_j$ is the coefficient on $x$ and $\mathbf{\Pi}_j^*$ is the coefficient on $z$. The rank condition we noted earlier is that what is estimated by $\mathbf{z}'\mathbf{f}/\mathbf{z}'\mathbf{z}$, which would correspond to $\mathbf{\Pi}_j^*$ not equal zero, meaning that it has rank 1.

Casual statements of the rank condition based on an IV regression of a variable $\mathbf{y}_{IV}$ on $(M_j + K_j)$ endogenous and exogeneous variables in $\mathbf{X}_{IV}$, using $K_j + K_j^*$ exogenous and instrumental variables in $\mathbf{Z}_{IV}$ (in the most familiar cases, $M_j = K_j^* = 1$), state that the rank requirement is that $(\mathbf{Z}_{\mathrm{IV}}'\mathbf{X}_{IV}/T)$ be nonsingular. In the notation we are using here, $\mathbf{Z}_{IV}$ would be $\mathbf{X} = (\mathbf{X}_j, \mathbf{X}_j^*)$ and $\mathbf{X}_{iv}$ would be $(\mathbf{X}_j, \mathbf{Y}_j)$. This nonsingularity would correspond to full rank of plim$(\mathbf{X}'\mathbf{X}/T)$ times plim$[(\mathbf{X}'\mathbf{X}^*/T, \mathbf{X}'\mathbf{Y}_j/T)]$ because plim$(\mathbf{X}'\mathbf{X}/T) = \mathbf{Q}$, which is nonsingular [see (10-40)]. The first $K_j$ columns of this matrix are the last $K_j$ columns of an identity matrix, which have rank $K_j$. The last $M_j$ columns are estimates of $\mathbf{Q}\overline{\mathbf{\Pi}}_j$, which we require to have rank $M_j$, so the requirement is that $\overline{\mathbf{\Pi}}_j$ have rank $M_j$. But, if $K_j^* \geq M_j$ (the order condition), then all that is needed is rank$(\mathbf{\Pi}_j^*) = M_j$, so, in practical terms, the casual statement is correct. It is stronger than necessary; the formal mathematical condition is only that the lower half of the matrix must have rank $M_j$, but the practical result is much easier to visualize.

It is also easy to verify that the rank condition requires that the predictions of $\mathbf{Y}_j$ using $(\mathbf{X}_j, \mathbf{X}_j^*)\overline{\mathbf{\Pi}}_j$ be linearly independent. Continuing this line of thinking, if we use 2SLS, the rank condition requires that the predicted values of the included endogenous variables not be collinear, which makes sense.

### 10.4.4 SINGLE EQUATION ESTIMATION AND INFERENCE

For purposes of estimation and inference, we write the model in the way that the researcher would typically formulate it,

$$\begin{aligned}\mathbf{y}_j &= \mathbf{X}_j\boldsymbol{\beta}_j + \mathbf{Y}_j\boldsymbol{\gamma}_j + \boldsymbol{\varepsilon}_j \\ &= \mathbf{Z}_j\boldsymbol{\delta}_j + \boldsymbol{\varepsilon}_j,\end{aligned} \tag{10-41}$$

where $\mathbf{y}_j$ is the "dependent variable" in the equation, $\mathbf{X}_j$ is the set of exogenous variables that appear in the $j$th equation—note that this is not all the variables in the model—and $\mathbf{Z}_j = (\mathbf{X}_j, \mathbf{Y}_j)$. The full set of exogenous variables in the model, including $\mathbf{X}_j$ and variables that appear elsewhere in the model (including a constant term if any equation includes one), is denoted $\mathbf{X}$. For example, in the supply/demand model in Example 10.6, the full set of exogenous variables is $\mathbf{X} = (\mathbf{1}, \mathbf{x}, \mathbf{z})$, while $\mathbf{X}_{Demand} = (\mathbf{1}, \mathbf{x})$ and $\mathbf{X}_{Supply} = (\mathbf{1}, \mathbf{z})$. Finally, $\mathbf{Y}_j$ is the endogenous variables that appear on the right-hand side of the $j$th equation. Once again, this is likely to be a subset of the endogenous variables in the full model. In Example 10.6, $\mathbf{Y}_j = (\mathbf{price})$ in both cases.

There are two approaches to estimation and inference for simultaneous equations models. **Limited information estimators** are constructed for each equation individually. The approach is analogous to estimation of the seemingly unrelated regressions model in Section 10.2 by least squares, one equation at a time. **Full information estimators** are used to estimate all equations simultaneously. The counterpart for the seemingly unrelated regressions model is the feasible generalized least squares estimator discussed in Section 10.2.3. The major difference to be accommodated at this point is the endogeneity of $\mathbf{Y}_j$ in (10-41).

The equation in (10-41) is precisely the model developed in Chapter 8. Least squares will generally be unsuitable as it is inconsistent due to the correlation between $\mathbf{Y}_j$ and $\boldsymbol{\varepsilon}_j$. The usual approach will be two-stage least squares as developed in Sections 8.3.2 through 8.3.4. The only difference between the case considered here and that in Chapter 8 is the source of the instrumental variables. In our general model in Chapter 8, the source of the instruments remained somewhat ambiguous; the overall rule was "outside the model." In this setting, the instruments come from elsewhere in the model—that is, "not in the $j$th equation." For estimating the linear simultaneous equations model, the most common estimator is

$$\begin{aligned}\hat{\boldsymbol{\delta}}_{j,\,2\,\mathrm{SLS}} &= [\hat{\mathbf{Z}}_j'\hat{\mathbf{Z}}_j]^{-1}\hat{\mathbf{Z}}_j'\mathbf{y}_j \\ &= [(\mathbf{Z}_j'\mathbf{X})(\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\mathbf{Z}_j)]^{-1}(\mathbf{Z}_j'\mathbf{X})(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}_j,\end{aligned} \tag{10-42}$$

where all columns of $\hat{\mathbf{Z}}_j'$ are obtained as predictions in a regression of the corresponding column of $\mathbf{Z}_j$ on $\mathbf{X}$. This equation also results in a useful simplification of the estimated asymptotic covariance matrix,

$$\text{Est.Asy.Var}[\hat{\boldsymbol{\delta}}_{j,\,2\,SLS}] = \hat{\sigma}_{jj}(\hat{\mathbf{Z}}_j'\hat{\mathbf{Z}}_j)^{-1}.$$

It is important to note that $\sigma_{jj}$ is estimated by

$$\hat{\sigma}_{jj} = \frac{(\mathbf{y}_j - \mathbf{Z}_j\hat{\boldsymbol{\delta}}_j)'(\mathbf{y}_j - \mathbf{Z}_j\hat{\boldsymbol{\delta}}_j)}{T}, \tag{10-43}$$

using the original data, not $\hat{\mathbf{Z}}_j$.

Note the role of the order condition for identification in the two-stage least squares estimator. Formally, the order condition requires that the number of exogenous variables that appear elsewhere in the model (not in this equation) be at least as large as the number of endogenous variables that appear in this equation. The implication will be that we are going to predict $\mathbf{Z}_j = (\mathbf{X}_j, \mathbf{Y}_j)$ using $\mathbf{X} = (\mathbf{X}_j, \mathbf{X}_j^*)$. In order for these predictions to be linearly independent, there must be at least as many variables used to compute the predictions as there are variables being predicted. Comparing $(\mathbf{X}_j, \mathbf{Y}_j)$ to $(\mathbf{X}_j, \mathbf{X}_j^*)$, we see that there must be at least as many variables in $\mathbf{X}_j^*$ as there are in $\mathbf{Y}_j$, which is the order condition. The practical rule of thumb that every equation have at least one variable in it that does not appear in any other equation will guarantee this outcome.

Two-stage least squares is used nearly universally in estimation of linear simultaneous equation models—for precisely the reasons outlined in Chapter 8. However, some applications (and some theoretical treatments) have suggested that the **limited information maximum likelihood (LIML) estimator** based on the normal distribution may have better properties. The technique has also found recent use in the analysis of weak instruments. A result that emerges from the derivation is that the LIML estimator has the same asymptotic distribution as the 2SLS estimator, and the latter does not rely on an assumption of normality. This raises the question why one would use the LIML technique given the availability of the more robust (and computationally simpler) alternative. Small sample results are sparse, but they would favor 2SLS as well.[25] One significant virtue of LIML is its invariance to the normalization of the equation. Consider an example in a system of equations,

$$y_1 = y_2\gamma_2 + y_3\gamma_3 + x_1\beta_1 + x_2\beta_2 + \varepsilon_1.$$

An equivalent equation would be

$$\begin{aligned} y_2 &= y_1(1/\gamma_2) + y_3(-\gamma_3/\gamma_2) + x_1(-\beta_1/\gamma_2) + x_2(-\beta_2/\gamma_2) + \varepsilon_1(-1/\gamma_2) \\ &= y_1\tilde{\gamma}_1 + y_3\tilde{\gamma}_3 + x_1\tilde{\beta}_1 + x_2\tilde{\beta}_2 + \tilde{\varepsilon}_1. \end{aligned}$$

The parameters of the second equation can be manipulated to produce those of the first. But, as you can easily verify, the 2SLS estimator is not invariant to the normalization of the equation—2SLS would produce numerically different answers. LIML would give the same numerical solutions to both estimation problems suggested earlier. A second virtue is LIML's better performance in the presence of weak instruments.

The LIML, or **least variance ratio** estimator, can be computed as follows.[26] Let

$$\mathbf{W}_j^0 = \mathbf{E}_j^{0\prime}\mathbf{E}_j^0, \tag{10-44}$$

where

$$\mathbf{Y}_j^0 = [\mathbf{y}_j, \mathbf{Y}_j],$$

and

$$\mathbf{E}_j^0 = \mathbf{M}_j\mathbf{Y}_j^0 = [\mathbf{I} - \mathbf{X}_j(\mathbf{X}_j'\mathbf{X}_j)^{-1}\mathbf{X}_j']\mathbf{Y}_j^0. \tag{10-45}$$

[25]See Phillips (1983).

[26]The LIML estimator was derived by Anderson and Rubin (1949, 1950). [See, also, Johnston (1984).] The much simpler and equally efficient two-stage least squares estimator remains the estimator of choice.

Each column of $\mathbf{E}_j^0$ is a set of least squares residuals in the regression of the corresponding column of $\mathbf{Y}_j^0$ on $\mathbf{X}_j$, that is, only the exogenous variables that appear in the $j$th equation. Thus, $\mathbf{W}_j^0$ is the matrix of sums of squares and cross products of these residuals. Define

$$\mathbf{W}_j^1 = \mathbf{E}_j^{1\prime}\mathbf{E}_j^1 = \mathbf{Y}_j^{0\prime}[\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{Y}_j^0. \tag{10-46}$$

That is, $\mathbf{W}_j^1$ is defined like $\mathbf{W}_j^0$ except that the regressions are on all the $x$'s in the model, not just the ones in the $j$th equation. Let

$$\lambda_1 = \text{smallest characteristic root of } (\mathbf{W}_j^1)^{-1}\mathbf{W}_j^0. \tag{10-47}$$

This matrix is asymmetric, but all its roots are real and greater than or equal to 1. [Depending on the available software, it may be more convenient to obtain the identical smallest root of the symmetric matrix $\mathbf{D} = (\mathbf{W}_j^1)^{-1/2}\mathbf{W}_j^0(\mathbf{W}_j^1)^{-1/2}$.] Now partition $\mathbf{W}_j^0$ into $\mathbf{W}_j^0 = \begin{bmatrix} w_{jj}^0 & \mathbf{w}_j^{0\prime} \\ \mathbf{w}_j^0 & \mathbf{W}_{jj}^0 \end{bmatrix}$ corresponding to $[\mathbf{y}_j, \mathbf{Y}_j]$, and partition $\mathbf{W}_j^1$ likewise. Then, with these parts in hand,

$$\hat{\boldsymbol{\gamma}}_{j,\,\text{LIML}} = [\mathbf{W}_{jj}^0 - \lambda_1\mathbf{W}_{jj}^1]^{-1}(\mathbf{w}_j^0 - \lambda_1\mathbf{w}_j^1) \tag{10-48}$$

and

$$\hat{\boldsymbol{\beta}}_{j,\,\text{LIML}} = (\mathbf{X}_j'\mathbf{X}_j)^{-1}\mathbf{X}_j'(\mathbf{y}_j - \mathbf{Y}_j\hat{\boldsymbol{\gamma}}_{j,\,\text{LIML}}).$$

Note that $\boldsymbol{\beta}_j$ is estimated by a simple least squares regression.[See (3-18).] The asymptotic covariance matrix for the LIML estimator is identical to that for the 2SLS estimator.

### *Example 10.8 Simultaneity in Health Production*

Example 7.1 analyzed the incomes of a subsample of Riphahn, Wambach, and Million's (2003) data on health outcomes in the German Socioeconomic Panel. Here we continue Example 10.4 and consider a Grossman (1972) style model for health and incomes. Our two-equation model is

$$\begin{aligned} \textit{Health Satisfaction} &= \alpha_1 + \gamma_1 \textit{ ln Income} + \alpha_2 \textit{ Female} + \alpha_3 \textit{ Working} + \alpha_4 \textit{ Public} + \alpha_5 \textit{ Add On} \\ &\quad + \alpha_6 \textit{ Age} + \varepsilon_H, \\ \textit{ln Income} &= \beta_1 + \gamma_2 \textit{ Health Satisfaction} + \beta_2 \textit{ Female} + \beta_3 \textit{ Education} + \beta_4 \textit{ Married} \\ &\quad + \beta_5 \textit{HHKids} + \beta_6 \textit{Age} + \varepsilon_I. \end{aligned}$$

For purposes of this application, we avoid panel data considerations by examining only the 1994 wave (cross section) of the data, which contains 3,377 observations. The health outcome variable is *Self Assessed Health Satisfaction* (HSAT). Whether this variable actually corresponds to a commonly defined objective measure of health outcomes is debateable. We will treat it as such. Second, the variable is a scale variable, coded in this data set 0 to 10. [In more recent versions of the GSOEP data, and in the British (BHPS) and Australian (HILDA) counterparts, it is coded 0 to 4.] We would ordinarily treat such a variable as a discrete ordered outcome, as we do in Examples 18.14 and 18.15. We will treat it as if it were continuous in this example, and recognize that there is likely to be some distortion in the measured effects that we are interested in. *Female*, *Working*, *Married*, and *HHkids* are dummy variables, the last indicating whether there are children living in the household. *Education* and *Age* are in years. *Public* and *AddOn* are dummy variables that indicate whether the individual takes up the public health insurance and, if so, whether he or she also takes up the additional